

口语对话系统中的语句主题提取

王 彬 江铭虎

(清华大学中文系计算语言研究室,北京 100084)

E-mail: wang-b01@mails.tsinghua.edu.cn

摘 要 语句的主题提取是口语对话系统中话语分析部分的工作。目前的口语对话系统大多将自然语言处理的重点放在语法和语义平面,而忽视了对上下文语境的分析,该文提出一种基于规则的语句主题提取方法,通过自底向上与自顶向下两种分析器完成主题与用户意图的提取,为系统的自然语言生成提供更准确的领域知识,从而大大提高了系统的整体性能。

关键词 主题提取 口语对话系统 自然语言理解

文章编号 1002-8331-(2004)18-0058-03 文献标识码 A 中图分类号 TP391

Extraction of the Topic from Utterances in a Spoken Dialog System

Wang Bin Jiang Minghu

(Computational Linguistics Lab, Chinese Department, Tsinghua University, Beijing 100084)

Abstract: Extraction of the topic from utterances is part of the discourse analysis in a spoken dialog system. At present, many spoken dialog systems take the key point on semantic surface, ignoring the context dependent interpretation of utterances. This paper presents an approach to extraction of user's intentions and topics through bottom-up and top-down analysis. It can provide more knowledge for system and therefore enhance the system performance greatly.

Keywords: topic extraction, spoken dialog system, natural language understanding

1 引言

人机口语对话系统是语音识别技术走向实用的一个重要研究方向。口语对话系统的目标是能够让人通过自然语言表达自己的思想,与计算机就某一领域的内容进行信息交互^[1]。近年来各国都投入了大量人力、物力、财力来研究口语对话系统,美国有 DARPA 的 Communicator 计划,欧洲有 ARISE 计划、REWARD 计划、VERBMOBIL 计划等。很多著名的学府与研究机构都在开展这项研究,如 MIT 的 SLS 实验室、CMU 的 ISL 实验室、Lucent-Bell 实验室、日本的 ATR 实验室、OGI 的 CSLU 中心和 Philips 公司等^[2]。国内也有中科院自动化所、清华大学、香港中文大学、台湾大学等多家研究单位从事此方面研究。

口语对话系统可分为四个层次:人机交互层、自然语言处理层、对话管理层、应用程序层。目前很多口语对话系统都将自然语言处理层研究的重点放在语法和语义平面,这样处理的一个问题是无法理清一段对话的整体内在联系^[3]。而对话往往由于口语中省略、指代、结构歧义等现象的存在,使得分析的结果具有歧义。这就要求我们用话语分析(Discourse Analysis)模块利用上下文语境和相关的领域知识进行排歧从而能得到最后的语义表示^[4]。话语分析后存储的对话历史还可以帮助系统推测用户下面将说的话,以实现语言处理模型的动态转换,从而提高系统识别的准确率。

话语分析包括两方面:一是从独立的对话中抽取主题和用户意图,二是用恰当的数据结构描述出主题与意图的转换关系^[5]。话语分析策略可以分为基于知识的方法与基于语料库的方法。基于知识的方法用一系列规则从对话中抽取主题和用户意图,并用规则的方法描述状态的转换过程。这些规则的设计

主要根据语言学者的总结^[6]。而基于语料库的策略需要用到两个概率: $P(T|W)$ 和 $P(I|W)$ 。 $P(T|W)$ 是主题 T 的条件概率, $P(I|W)$ 是用户意图 I 在一个对话中出现过的词符集 W 下的条件概率。这两个概率通过对已标注的语料库的分析来估测,并用来抽取主题和识别用户意图。 n 元语法模型或隐马尔可夫模型(HMM)也可以用来描述对话中的主题与用户意图的转移过程^[7]。语料库的规模越大,标注越精准,建立起来的对话模型也就越可靠。

论文提出了一种基于知识的主题抽取方法,以实现从一个口语对话系统的用户话语中抽取该句的主题和用户意图。该口语对话系统的主要任务是导购信息服务。作者用两种知识源来描述对话历史:一个是用户意图转移网络图,用来描述从一个意图到另一个意图的可能转换途径;另一个是主题树,由查询对话中可能出现的主题构成。假设主题是沿着主题树发展的。这样当对话进行时,不同主题的发展就会形成一棵子树,文中称之为动态主题树。

用自底向上和自顶向下两种分析方法来提取主题。自底向上的方法是通过将一系列预先设定的规则应用于话语的语义解释中,从而得到主题与用户意图的候选者。自顶向下的方法主要是由对话历史的当前状态决定候选者,而对话历史则由对用户意图转移网络图的追踪回溯和动态主题树来反映。将两种方法得到的候选者通过逻辑“与”的操作最终得到用户意图和话语的主题。

2 主题提取

2.1 话语分析部分的总体框架

图 1 所示是话语分析部分的整体流程。

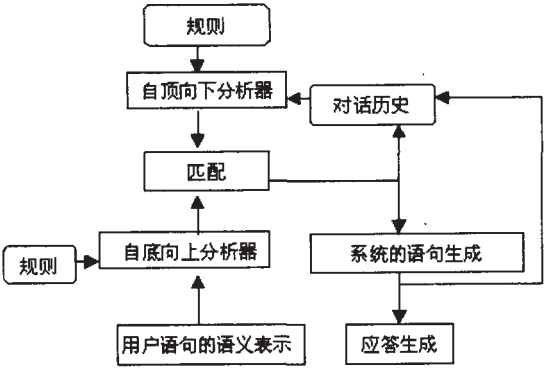


图 1 话语分析的总体框图

在口语对话系统中话语分析部分处于对话管理层的位置，直接与对话管理打交道。它所面临的处理对象是已经过句法分析和语义分析后的句子。本系统中该部分包括自底向上和自顶向下两个分析器：自底向上分析器产生主题和用户意图的自底向上候选结果；自顶向下分析器通过参考对话历史来预测对话中可能产生的下一句子的主题和用户意图。对话历史由与或主题树和用户意图转移网络状态图表示。自顶向下分析器得到的候选者将以启发式算法排序。

自底向上得到的结果将与自顶向下得到的预测去匹配，最佳匹配结果就是该句子的主题和用户意图，并被保留记录在对话历史中，传递到系统的语句生成模块。对话系统所生成的下一语句的主题和意图将由这个模块决定，结果也要记录至对话历史中。

系统的语义分析模块是基于格语法的。在格语法中，一句话的语义由句中主要动词所对应的格框架表示。一个格框架由一系列槽构成，每个槽代表某类词与这个主要动词之间的关系。譬如一个名词短语与动词可以构成施事、工具等语义关系。一句话中的名词短语将根据名词短语的语义标记去填充该句中动词所形成的框架槽。这样，一句话语的语义表示就由下面三项来表示：一个主动词，一个填充了语义槽的格框架和该句话所对应的形态信息。

2.2 知识源表示

口语对话系统中话语分析部分的知识源主要是对话历史，而对话历史的表示依据于主题和用户意图的组织。这里利用主题树和用户意图转移网络图来描述对话历史。

S39 您好，这里是北京地区购物查询系统 Gogo。请问您需要什么？
U39 我想买电热水器。
S40 请问您想要国产电热水器还是进口电热水器？
U40 最好是国产的吧。
S41 国产电热水器有万家乐、海尔、阿里斯顿等。
U41 那么在哪里可以买到呢？
S42 各大商场和电器城。
U42 电器城都有哪些呢？
S43 北京地区的电器城有国美电器、苏宁电器等。
U43 国美电器怎么走啊？
S44 您是问北京国美电器城的地址吗？
U44 是啊。
S45 国美北太平庄商城地址是北太平庄路口 300 米路南，咨询电话 62071730。
U45 好，谢谢了。
S46 很高兴能为您服务。

图 2 对话示例

中系统先向用户问候，明确用户所需服务的类型：购买。然后确定所购商品的要求，并提供了几个牌子的电器作为候选项。用户继续询问购物地点的相关细节，系统再次给出候选项，并最终给出一个商场的具体位置。用户以感谢作为结束语，系统也给予相应的回答。在这样一种目标确定的对话系统中，主题是沿着任务相关树的结构转移的。例如图 2 中所给示例的主题变化过程可由图 3 所示。

考察了对话中主题变化的特点，用一个“主题框架”集合作为主题的知识源。一个主题框架将相关的主题归入在查询对话系统中可能出现的同一框架中，譬如表示“家电”的框架可以包括家电品牌、销售地点、促销活动等等。而一个框架的槽值可以是其他框架，这样这些主题框架就形成了一棵树。假设主题的发展是沿着主题树的，当对话进行发展时就会形成前面提到过的动态主题树。

主题的活动变化可以用一棵与或树来表示。在一棵与或树中，“与”结点代表由用户引入的主题，“或”结点代表由系统引入的主题。譬如当用户要查询两个或更多的商品信息时，每个商品代表一个主题，因此系统必须提供所有这些主题的信息；而如果系统为用户提出两个或更多的候选项时，用户无需对这些候选项都做出反应，可以只对其中一个发生兴趣从而提出问题，使对话的语句转移到下一个主题。因此用与或树来反应用户与系统引入主题的差异性是比较合适的。

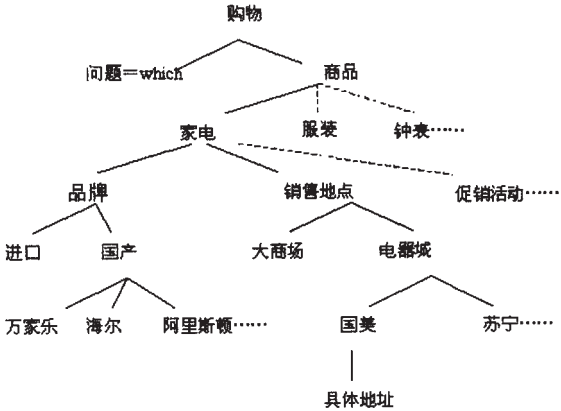


图 3 主题表示例

动态主题树的结点也可以分为原型结点和例结点两种。原型结点直接从主题树继承而来，而例结点是由数据库中的数据构成。例结点是动态主题树的叶结点，因此不能再被扩展。一个动态主题树的结点也可分为下列四种状态：焦点态(focused)、悬挂态(suspended)、结束态(closed)和未访问态(unvisited)。这些状态信息可用于焦点转移控制。

一段对话中每个句子的表达都有一定的目的，也就是说用户想传达给对话系统的意图，这就是所说的用户意图(user intention)。需要声明的是这里所说的用户意图是一广义的概念，它不仅指用户所说句子表达的意图，有时也包括系统生成句子所表达的意图。表 1 列举了与图 1 中出现的句子相对应的用户意图类型。确认意图既可以是用户确认他的知识源，也可以是系统确认它的反应，这样可以帮助系统澄清口语对话中出现的犹豫现象、改正一些语音识别错误。表 1 中所列举的用户意图是作者所组织的用户意图中的最高层，每一个用户意图还可以再根据相关领域细分为更多的子范畴。本口语对话系统有一个

用户意图转移网络图作为知识源 图中的每一个状态都代表一个用户意图。这样这个网络图就可以描述对话中用户意图转移的过程。因此对话历史就可以由图中的某些状态来表示。

表 1 用户意图

用户意图	示例
GR 问候(Greeting)	S39
AI 询问信息(Ask-information)	S40 ,U41 ,U42 ,U43 ,
GI 给定信息(Give-information)	U41 ,S42 ,S43 ,S45
CN 确认(Confirmation)	S44
RC 确认反应(Response-to-confirmation)	U44
ED 结束(ending)	S45

2.3 主题及意图的提取

本口语对话系统中接受的词汇根据语义聚类并组织成树结构。上层结点(靠近根部)所对应的词汇被称为概念词汇,对句子的阐释起到语义标注的作用。第二部分中提到,互相关联的主题形成一个主题框架。一个主题框架中包含若干槽,这些槽被句中相应的词语填入,由于每个词语都有语义标注,这个槽也就具有了语义标注信息。在进行自底向上分析时会用到主题、槽和槽所对应的语义标注之间的关系。将这个关系组织成一张主题-槽表,也就是说可以通过查表中的单词得到一个主题与一个槽之间的关系。

另外还用到一张由动词槽框架得到的动词-焦点表。在所设定的购物查询这一受限领域,假设一句话中的某一个词填充到该句主动词的槽中,且该单词带有特定语义标记,则这一单词很可能成为被关注的主题。这张动词-焦点表就是要描述这种关系。

自底向上分析器通过这两张表得到主题的候选者。对于一个待分析语句中的每一个格框架槽的填充词,首先通过查找主题-槽表找出一对主题和语义槽,然后具有相同主题的主题-语义槽对将被归并到有两个或更多填满语义槽的主题框架中。最后余下的主题框架将被剪枝掉,通过查询动词-焦点表来决定焦点槽。

而关于主题的自顶向下分析器的工作有赖于对话历史、动态主题树和焦点结点、系统预测出的该句用户意图等多方面知识。这里要用到焦点-转移规则来进行自顶向下分析。这些规则描述了动态主题树中的结点如何根据焦点的转换而移动,并根据每次焦点移动可能发生的程度进行排序。譬如,当某一句话的用户意图属于“询问信息”类型时,被问部分就将被关注,而用户应该去回答这个问题,这就意味着焦点没有移动。而当系统生成的某句话中的意图属于“给定信息”时,用户将获得他所提问题的答案,也就是说该主题已经被关注完,因此焦点可以转移到对话所能保持范围内的任一主题上。将焦点-转移规则应用在动态主题树当前被关注的节点上,根据主题的优先级就可以预测出用户下一句话中将出现的主题。

本对话系统为用户意图的自底向上分析器维护了一张表,用来描述一句话中表示意图的动词和形态信息的关系。在这张表中,表示意图的动词被分为三种类型(1)用来表示用户行为的动词,如“买”、“走”(2)表示用户要求的动词,如“告诉我”;(3)表示用户反应的动词,如“知道了”、“好”。而按形式又可以把意图分为问题与非问题两种。表 1 中的用户意图按形式分类后如表 2 所示。

而反应用户意图的对话历史将由状态转移网络图中的当前状态反应。这张表描述用户意图的可能转移情况。而转移图

中由某状态出发可以达到的下一状态就代表了对用户意图预测的结果。

表 2 用户意图的分类

预测类型	形态	
	问题	非问题
(1)	AI ,CN	GI
(2)	AI ,CN	AI
(3)		AI ,GI ,RC ,AK

3 实验结果

为了测试上面介绍的主题与意图自动抽取方法,作者从一个包含约 9 万语句的对话语料库中整理出与导购任务有关的对话,然后从中随机抽取了一些对话及其包括的句子,并通过网络搜索和实际商场考察收集了一些在进行购物查询时经常会发生的对话,组成了测试语料库。

作者将上述方法应用到用户话语的测试中。表 3 显示了只用自底向上分析器时提取的准确率和把自底向上与自顶向下结合起来后的测试结果。UQ 一列表示当分析器得到的句子的主题或意图候选者只有一个时抽取的准确率,ML 一列表示当有多个候选结果且可匹配到一起时的抽取准确率,ER 一列表示当无互相匹配的候选结果或没有产生候选项时的抽取结果。

由图中看到,当只用自底向上分析器来进行抽取时,主题和用户意图的准确率分别是 63.1%和 71.4%。而当自底向上与自顶向下分析器结合起来后,抽取的准确率有所提高,主题和用户意图对应的数字分别为 83.8%和 81.3%。虽然测试语料库规模较小,但这个结果仍可说明基于规则的主题提取方法达到了一定的效果。之所以在两种方法结合后仍有多多个候选结果的情况产生,是因为在排序过程中有可能产生具有相同优先级的自顶向下的候选者。而 ER 一列的数字在两种方法结合后有所增加是因为自顶向下的候选者与自底向上的候选者的交集有可能为空。

表 3 主题和用户意图提取结果

	自底向上(%)			结合后(%)		
	UQ	ML	ER	UQ	ML	ER
主题	63.1	28.4	8.5	83.8	7.1	9.1
用户意图	71.4	21.3	7.3	81.3	10.1	8.6

UQ :只有一个候选结果时
ML :有多多个候选结果时
ER :无互相匹配的候选结果或没有产生候选项

4 结论

该文介绍了一种基于知识的口语对话系统中语句主题和意图的自动提取方法。这个方法包括自底向上分析器和自顶向下分析器两个部分。自底向上分析器通过将设定的规则应用于句子语义输入从而得到该句主题和意图的候选项,而自顶向下分析器通过参照对话历史得到该句的主题和意图候选项。然后将两种方法得到的结果进行匹配,此时得到的最佳匹配被认为是该句的主题和意图。

作者用一个具有一定规模的语料库来检测此种方法,结果是主题提取的准确率达到 83.8%,用户意图提取的准确率达到 81.3%。这说明此方法对提高系统的性能有很大帮助。作者进一步的工作设想是补充设计更多的提取规则,扩大主题覆盖面,

(下转 98 页)

训练集 4000 个作为测试集。输入维数 16 ,输入值是 0 到 15 的整数。Holland-style adaptive classifier systems 是第一个做这个样本集的方法 ,达到的最好测试集分类正确率是 82.7%左右。

训练 : 首先将训练集中的 16000 个样本按类排列。按 4.1 中的方法 ,每个类生成了 60 个左右的核函数(即两层非线性网络的输出节点) ,再将两层非线性网络的输出作为两层线性网络的输入 ,训练两层线性网络的权值。26 个类生成了 26 个如图 2 所示的串联网。测试 :如实验 1 所述。

NB 算法的正确率最低 ,仅为 74.9%。IB1 的正确率最高 ,达到 95.7%。该文方法测试集的正确率为 87% ,比 NB 算法高 12% ,比多层感知器高 7.7%。训练集的正确率达到 96.8% ,这说明该文方法有很强的适应能力 ,这与算法的自动确定中心数目和自适应调整核函数宽度的特点有关。另外 ,由于我们采用的是局部训练的策略 ,每一个类单独生成一个串联网 ,这些网络可以并行训练 ,每个网络在 PC 机上平均的训练时间大约为 1.5 小时 ,总的训练时间比那些全局训练的网络要少得多 ,且不会产生不收敛的情况 ,这对大样本数据的训练是有益的。测试集的分类正确率也较高 ,说明网络的泛化能力较强。我们发现 ,在训练时适当放宽“子类增长结束条件” ,使子类中包含几个其它类的样本 ,就是让这些样本点包含在它的椭圆之内 ,训练集的分类正确率有一定程度的降低 ,但测试集的分类正确率有所提高。另外 ,子类之间的相互重叠 ,也就是一个样本可属于多个子类 ,将测试集的分类正确率提高了 8%左右 ,这是由于重叠使子类所能覆盖的范围变大了 ,因而泛化能力增强了。

表 1 各种方法的分类正确率比较

方法	测试集(%)
NB(Naive Bayesian classifier)	74.9
K-最近邻算法	89.9
IB1(a minor variation of k-NN)	95.7
C4.5(决策树)	77.7
多层感知器	79.3
该文方法	87

6 总结

该文在大样本数据识别的应用背景下 ,提出了对 RBF 网络拓扑结构及核函数生成算法的改进。将前向三层 RBF 网络改为一个两层非线性网络和一个两层线性网络的串联 ,简化了核函数参数的训练 ,并提高了训练速度 ,而且这种网络结构的改进是该文核函数生成算法的基础 ,因为后者是根据核函数值的误差来修正 μ 和 σ 的。该文的核函数生成算法能够自动确定核函数数目 ,并且能自适应调整核函数中心和宽度 ,避免了

人为确定核函数数目和宽度所带来的盲目性 ,增强了网络对不同样本集的适应能力。另外根据该文算法每个类对应一个网络 ,这些网络相互没有影响 ,可以分开并行训练 ,减小了时间复杂度和空间复杂度 ,有利于大样本数据的识别。

该文方法的泛化能力还可加强 ,可以通过对参数 A、B 以及“子类增长结束条件”的调整 ,增加子类中包含的样本数 ,另外还要研究如何处理“孤岛”(即只包含一个样本的子类) 。对高斯函数的形式还可以作修改 ,使其对不同的样本集有更强的适应能力。(收稿日期 2003 年 8 月)

参考文献

1.Gao Daqi ,Wushouyi.An optimization method for the topological structures of feed-forward multi-layer neural networks[J].Pattern Recognition ,1998 ;31(9) :1337~1342
2.Gao Daqi ,Yanggenxing.Basic Principles of Pattern Classification Methods Based on Improved RBF Neural Networks[J].Journal of East China University of Science and Technology ,2001 ;27(6) :677~683
3.Karayiannis N B ,Mi G W.Growing radial basis neural networks : merging supervised and unsupervised learning with network growth technique[J].IEEE Transactions on Neural Networks ,1997 ;8(6) :1492~1560
4.Roy A ,Govil S ,Miranda P.An algorithm to generate radial basis function(RBF)-like nets for classification problem[J].Neural Networks , 1995 ;8(2) :179~201
5.Karayusnnis N B.Reformulated radial basis neural networks trained by gradient descent[J].IEEE Trans on Neural Networks ,1999 ;10(3) : 657~671
6.Murphy P M ,Aha P.The UCI repository of machine learning data-bases and domain theories.http ://www.ics.uci.edu/~mlearn ,1995
7.Friedhelm Schwenker ,Hans A Kestler ,Gunther Palm.Three learning phases for radial-basis-function networks[J].Neural Networks 2001 ;14 : 439~458
8.Musavi M T ,Ahmed W ,Chan K H et al.On the training of radial basis function classifier[J].Neural Networks ,1992 ;5(2) :595~603
9.Zhu Q ,Cai Y ,Liu L.A global learning algorithm for a RBF network[J]. Neural Networks ,1999 ;12(2) :527~540
10.Uykan Z ,Guzelis C ,Celebi E et al.Analysis of input-output clustering for determining centers of RBFN[J].IEEE transactions on Neural Networks 2000 ;11(4) :851~857
11.A Esposito ,M Marinaro ,D Oricchio et al.Approximation of continous and discontinuous mappings by a growing neural RBF-based algorithm[J].Neural Networks 2000 ;13 :651~665
12.ftp ://ftp.ics.uci.edu/pub/mechine-learning-databases or http ://www. ics.uci.edu/~mlearn

(上接 60 页)

以得到更精确的结果。(收稿日期 2003 年 10 月)

参考文献

1.Huang Xuedong.Spoken Language Processing[M].New Jersey :Prentice Hall PTR 2001
2.Bo Zhang ,Qingsheng Cai ,Jianfeng Mao.Spoken Dialogue Management as Planning and Acting under Uncertainty[C].In :European Conference on Speech Communication and Technology(EuroSpeech)2001 ;3 : 2169~2172
3.Egbert Ammicht ,Alexandros Potamianos ,Eric Fosler-Lussier.Ambiguity Representation and Resolution in Spoken Dialogue Systems[C].In :EuroSpeech 2001 ;3 :2217~2220
4.Yunbiao XU ,Masahiro ARAKI ,Yasuhisa NIIMI.A Multilingual-supporting dialog System Using a Common dialog Controller[C].In :EuroSpeech 2001 ;2 :1283~1286
5.Y Niimi ,N Takinaga ,T Nishimoto.Extraction of the dialog act and the topic from utterances in a spoken dialog system[C].In :International Conference on Spoken Language Processing(ICSLP) ,1998 : 2079~2082
6.Oliver Lemon ,Anne Bracy ,Alexander Gruenstein ,Stanley Peters.The WITAS Multi-Modal Dialogue System II[C].In :EuroSpeech 2001 ;3 : 1559~1562
7.Konrad Scheffler ,Steve Young.Probabilistic Simulation of Human-Machine Dialogues[C].In :International Conference on Acoustics Speech and Signal Processing(ICASSP)2000 ,Istanbul ,2000 :1217~1220

4.Yunbiao XU ,Masahiro ARAKI ,Yasuhisa NIIMI.A Multilingual-supporting dialog System Using a Common dialog Controller[C].In :EuroSpeech 2001 ;2 :1283~1286
5.Y Niimi ,N Takinaga ,T Nishimoto.Extraction of the dialog act and the topic from utterances in a spoken dialog system[C].In :International Conference on Spoken Language Processing(ICSLP) ,1998 : 2079~2082
6.Oliver Lemon ,Anne Bracy ,Alexander Gruenstein ,Stanley Peters.The WITAS Multi-Modal Dialogue System II[C].In :EuroSpeech 2001 ;3 : 1559~1562
7.Konrad Scheffler ,Steve Young.Probabilistic Simulation of Human-Machine Dialogues[C].In :International Conference on Acoustics Speech and Signal Processing(ICASSP)2000 ,Istanbul ,2000 :1217~1220